

## SEQUENCE LISTING

PCT/PTO 22 NOV 2004

#4

<110> Dimitrov, Dimiter Stanchev  
Chow, Yen-Hung  
Phogat, Sanjay Kumar  
Broder, Christopher Charles  
The Government of the United States of America  
as represented by The Secretary of the  
Department of Health and Human Services

<120> HIV-1 Envelope Glycoproteins Stabilized by Flexible  
Linkers as Potent Entry Inhibitors and Immunogens

<130> 015280-458000US

<140> US 10/506,651

<141> 2004-09-02

<150> WO PCT/US02/07144

<151> 2002-03-05

<160> 33

<170> PatentIn Ver. 2.1

<210> 1

<211> 2562

<212> DNA

<213> Human immunodeficiency virus type 1

<220>

<223> full length human immunodeficiency virus-1 (HIV-1)  
envelope glycoprotein (Env, gp160) from primary  
R5X4 isolate 89.6

<400> 1

atgagagtga	aggagatcag	gaagaattgg	cagcacttga	gaggggggcat	cttgctcctt	60
gggatgttga	tgatctgtag	tgctgcaaaa	gaaaaaacgt	gggtcacaaat	ctattatggg	120
gtacctgtgt	ggagagaagc	aaccaccact	ctattttgtg	catcagatgc	taaagcctat	180
gatacagagg	tacataatgt	ttggggccaca	catgcctgtg	taccacacaga	ccccaaccca	240
caagaagtag	tattgggaaa	tgtgacagaa	aattttaaca	tgtggaaaaa	taacatggta	300
gatcagatgc	atgaggatat	aatcagttta	tgggatgaaa	gcctaaagcc	atgtgtaaaa	360
ttaacccccac	tctgtgttac	tttaaattgc	actaatttga	atatcactaa	gaatactact	420
aatcccacta	gtagcagctg	gggaatgatg	gagaaaggag	aaataaaaaa	ttgctctttc	480
tatatcacca	caagcataag	aaataaggta	aagaaagaat	atgcactttt	taatagactt	540
gatgtagtac	caatagaaaa	tactaataat	actaagtata	ggttaataag	ttgtaacacc	600
tcagtcatta	cacaggcctg	tccaaaggta	tccttttcagc	caattcccat	acattattgt	660
gtcccggctg	ggttttgcgat	gctaaagtgt	aacaataaga	cattcaatgg	atcaggacca	720
tgacaaaatg	tcagcacagt	acaatgtaca	catggaatta	ggccagtggg	gtcaactcaa	780
ctgctgttaa	atggcagctc	agcagaagaa	gacatagtaa	ttagatctga	aaatttcaca	840
gacaatgcta	aaaccataat	agtacagcta	aatgaatctg	tagtaattaa	ttgtacaaga	900
cccaacaaca	atacaagaag	aaggttatct	ataggaccag	ggagagcatt	ttatgcaaga	960
agaaacataa	taggagatat	aagacaagca	cattgtaaca	ttagtagagc	aaaatggaat	1020
aacacttttac	aacagatagt	tataaaatta	agagaaaaat	ttaggaataa	aacaatagcc	1080
tttaatcaat	cctcaggagg	ggacccagaa	attgtaatgc	acagttttta	ttgtggaggg	1140
gaattttttct	actgtaatac	agcacaactg	tttaatagta	cttggaatgt	tactggaggg	1200
acaaatggca	ctgaaggaaa	tgacataatc	acactccaat	gcagaataaa	acaaattata	1260
aatatgtggc	agaaagtagg	aaaagcaatg	tatgcccctc	ccatcacagg	acaaattaga	1320
tgttcataca	atattacagg	gctgctacta	acaagagatg	gaggtaatag	tactgagact	1380
gagactgaga	tcttcagacc	tggaggagga	gatatgaggg	acaattggag	aagtgaatta	1440
tataaatata	aagtagtaag	aattgaacca	ataggagtag	caccacaccg	ggcaaagaga	1500

```

agaacagtgc aaagagaaaa aagagcagtg ggaataggag ctgtgttcct tgggttcttg 1560
ggagcagcag gaagcactat gggcgagcg tcagtgcgc tgacgggtaca ggccaggcta 1620
ttattgtctg gtatagtgc gcagcagaac aatctgctga gggctattga ggcgcaacag 1680
catatgttgc aactcacagt ctggggcatc aagcagctcc aggcaagagt cctggctctg 1740
gaaagatacc taagggatca acagctcatg ggaatttggg gttgctctgg aaaactcatt 1800
tgcaccactt ctgtgccttg gaatgttagt tggagtaata aatctgtgga tgatatttgg 1860
aataacatga cctggatgga gtgggaaaga gaaattgaca attacacaga ctatatatat 1920
gacttacttg aaaaatcgca aaccacaaca gaaaagaatg aaaaagaatt attggaattg 1980
gataaatggg caagtttggtg gaattgggtt gacataacaa actggctgtg gtatataaga 2040
ttattcataa tgatagtagg aggccttgata ggtttaagaa tagtttttgc tgtactttct 2100
atagtaaata gagttaggca gggatattca ccattatcgt ttcagaccct cctcccagcc 2160
tcgaggggac ccgacaggcc cgaaggaaca gaagaagaag gtggagagag agacagagac 2220
agatccggtc cattagtga cggattcttg gcacttttct gggctcgattt gaggaacctg 2280
tgctcttcc tctaccacct cttgagaaac ttactcttga ttgtaacgag gattgtggaa 2340
cttctgggac gcaggggggtg ggaagccctc aaatattggt ggaatctcct gcaatattgg 2400
agccaggaac taaagaatag tgctgttagc ttgctcaacg ccacagccat agcagtagct 2460
gaggggacag atagggttat aaaaatagta caaagagctt gtagagctat tcgcaacata 2520
cctacaagaa tcagacaggg cttggaaagg gctttgctat aa 2562

```

```

<210> 2
<211> 853
<212> PRT
<213> Human immunodeficiency virus type 1

```

```

<220>
<223> full length human immunodeficiency virus-1 (HIV-1)
      envelope glycoprotein (Env, gp160) from primary
      R5X4 isolate 89.6

```

```

<400> 2
Met Arg Val Lys Glu Ile Arg Lys Asn Trp Gln His Leu Arg Gly Gly
  1              5              10              15

Ile Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Lys Glu Lys
  20              25              30

Thr Trp Val Thr Ile Tyr Tyr Gly Val Pro Val Trp Arg Glu Ala Thr
  35              40              45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
  50              55              60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
  65              70              75              80

Gln Glu Val Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys
  85              90              95

Asn Asn Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
  100             105             110

Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
  115             120             125

Asn Cys Thr Asn Leu Asn Ile Thr Lys Asn Thr Thr Asn Pro Thr Ser
  130             135             140

Ser Ser Trp Gly Met Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe
  145             150             155             160

```

Tyr Ile Thr Thr Ser Ile Arg Asn Lys Val Lys Lys Glu Tyr Ala Leu  
 165 170 175  
 Phe Asn Arg Leu Asp Val Val Pro Ile Glu Asn Thr Asn Asn Thr Lys  
 180 185 190  
 Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro  
 195 200 205  
 Lys Val Ser Phe Gln Pro Ile Pro Ile His Tyr Cys Val Pro Ala Gly  
 210 215 220  
 Phe Ala Met Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Ser Gly Pro  
 225 230 235 240  
 Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val  
 245 250 255  
 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Asp Ile  
 260 265 270  
 Val Ile Arg Ser Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val  
 275 280 285  
 Gln Leu Asn Glu Ser Val Val Ile Asn Cys Thr Arg Pro Asn Asn Asn  
 290 295 300  
 Thr Arg Arg Arg Leu Ser Ile Gly Pro Gly Arg Ala Phe Tyr Ala Arg  
 305 310 315 320  
 Arg Asn Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg  
 325 330 335  
 Ala Lys Trp Asn Asn Thr Leu Gln Gln Ile Val Ile Lys Leu Arg Glu  
 340 345 350  
 Lys Phe Arg Asn Lys Thr Ile Ala Phe Asn Gln Ser Ser Gly Gly Asp  
 355 360 365  
 Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr  
 370 375 380  
 Cys Asn Thr Ala Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Gly Gly  
 385 390 395 400  
 Thr Asn Gly Thr Glu Gly Asn Asp Ile Ile Thr Leu Gln Cys Arg Ile  
 405 410 415  
 Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys Ala Met Tyr Ala  
 420 425 430  
 Pro Pro Ile Thr Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu  
 435 440 445  
 Leu Leu Thr Arg Asp Gly Gly Asn Ser Thr Glu Thr Glu Thr Glu Ile  
 450 455 460  
 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu  
 465 470 475 480

Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Ile	Gly	Val	Ala	Pro	Thr	
				485					490					495		
Arg	Ala	Lys	Arg	Arg	Thr	Val	Gln	Arg	Glu	Lys	Arg	Ala	Val	Gly	Ile	
			500					505					510			
Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	
		515					520					525				
Ala	Ala	Ser	Val	Thr	Leu	Thr	Val	Gln	Ala	Arg	Leu	Leu	Leu	Ser	Gly	
		530				535					540					
Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	
545					550					555					560	
His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	
			565					570						575		
Val	Leu	Ala	Leu	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Met	Gly	Ile	
			580					585					590			
Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ser	Val	Pro	Trp	Asn	
		595					600						605			
Val	Ser	Trp	Ser	Asn	Lys	Ser	Val	Asp	Asp	Ile	Trp	Asn	Asn	Met	Thr	
	610					615						620				
Trp	Met	Glu	Trp	Glu	Arg	Glu	Ile	Asp	Asn	Tyr	Thr	Asp	Tyr	Ile	Tyr	
625					630					635					640	
Asp	Leu	Leu	Glu	Lys	Ser	Gln	Thr	Gln	Gln	Glu	Lys	Asn	Glu	Lys	Glu	
			645					650						655		
Leu	Leu	Glu	Leu	Asp	Lys	Trp	Ala	Ser	Leu	Trp	Asn	Trp	Phe	Asp	Ile	
			660					665						670		
Thr	Asn	Trp	Leu	Trp	Tyr	Ile	Arg	Leu	Phe	Ile	Met	Ile	Val	Gly	Gly	
		675					680					685				
Leu	Ile	Gly	Leu	Arg	Ile	Val	Phe	Ala	Val	Leu	Ser	Ile	Val	Asn	Arg	
	690					695					700					
Val	Arg	Gln	Gly	Tyr	Ser	Pro	Leu	Ser	Phe	Gln	Thr	Leu	Leu	Pro	Ala	
705					710					715					720	
Ser	Arg	Gly	Pro	Asp	Arg	Pro	Glu	Gly	Thr	Glu	Glu	Glu	Gly	Gly	Glu	
				725					730					735		
Arg	Asp	Arg	Asp	Arg	Ser	Gly	Pro	Leu	Val	Asn	Gly	Phe	Leu	Ala	Leu	
			740					745						750		
Phe	Trp	Val	Asp	Leu	Arg	Asn	Leu	Cys	Leu	Phe	Leu	Tyr	His	Leu	Leu	
		755					760					765				
Arg	Asn	Leu	Leu	Leu	Ile	Val	Thr	Arg	Ile	Val	Glu	Leu	Leu	Gly	Arg	
	770					775					780					
Arg	Gly	Trp	Glu	Ala	Leu	Lys	Tyr	Trp	Trp	Asn	Leu	Leu	Gln	Tyr	Trp	
785					790					795					800	

Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn Ala Thr Ala  
805 810 815

Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Lys Ile Val Gln Arg  
820 825 830

Ala Cys Arg Ala Ile Arg Asn Ile Pro Thr Arg Ile Arg Gln Gly Leu  
835 840 845

Glu Arg Ala Leu Leu  
850

<210> 3  
<211> 2051  
<212> DNA  
<213> Human immunodeficiency virus type 1

<220>  
<223> human immunodeficiency virus-1 (HIV-1) envelope  
glycoprotein gp140 truncated version of gp 160  
from primary R5X4 isolate 89.6

<400> 3  
atgagagtga aggagatcag gaagaattgg cagcacttga gagggggcat cttgctcctt 60  
gggatgttga tgatctgtag tgctgcaaaa gaaaaaacgt gggtcacaat ctattatggg 120  
gtacctgtgt ggagagaagc aaccaccact ctatcttctg catcagatgc taaagcctat 180  
gatacagagg tacataatgt ttgggccaca catgcctgtg taccacacaga cccaaccca 240  
caagaagtag tattgggaaa tgtgacagaa aattttaaca tgtggaaaaa taacatggta 300  
gatcagatgc atgaggatat aatcagttta tgggatgaaa gcctaaagcc atgtgtaaaa 360  
ttaacccact tctgtgttac tttaaattgc actaatttga atatcactaa gaatactact 420  
aatcccaact gtagcagctg gggaatgatg gagaaaggag aaataaaaaa ttgctctttc 480  
tatatcacca caagcataag aaataaggta aagaaagaat atgcactttt taatagactt 540  
gatgtagtac caatagaaaa tactaataat actaagtata ggtaataaag ttgtaacacc 600  
tcagtcatta cacaggcctg tccaaaggta tcctttcagc caattcccat acattattgt 660  
gtcccggctg ggtttgcatg gctaaagtgt aacaataaga cattcaatgg atcaggacca 720  
tgcacaaatg tcagcacagt acaatgtaca catggaatta ggccagtggg gtcaactcaa 780  
ctgctgttaa atggcagttc agcagaagaa gacatagtaa ttagatctga aaatttcaca 840  
gacaatgcta aaaccataat agtacagcta aatgaatctg tagtaattaa ttgtacaaga 900  
cccaacaaca atacaagaag aaggttatct ataggaccag ggagagcatt ttatgcaaga 960  
agaacataaa taggagatat aagacaagca cattgttaaca ttagtagagc aaaatggaat 1020  
aacactttac aacagatagt tataaaatta agagaaaaat ttaggaataa aacaatagcc 1080  
tttaatcaat cctcaggagg ggaccagaa attgtaatgc acagttttta ttgtggaggg 1140  
gaatttttct actgtaatac agcacaactg tttaatagta cttggaatgt tactggaggg 1200  
acaaatggca ctgaaggaaa tgacataatc aactccaat gcagaataaa acaaattata 1260  
aatatgtggc agaaagtagg aaaagcaatg tatgcccctc ccatcacagg acaaattaga 1320  
tggtcatcaa atattacagg gctgctacta acaagagatg gaggtaatat tactgagact 1380  
gagactgaga tcttcagacc tggaggagga gatatgaggg acaattggag aagtgaatta 1440  
tataaatata aagtagtaag aattgaacca ataggagtag caccaccag ggcaagaga 1500  
agaacagtgc aaagagaaaa aagagcagtg ggaataggag ctgtgttcct tgggttcttg 1560  
ggagcagcag gaagcactat gggcgagcg tcagtgcgcg tgacgggtaca ggccaggcta 1620  
ttattgtctg gtatagtgcg gcagcagaac aatctgctga gggctattga ggcgcaacag 1680  
catatgttgc aactcacagt ctggggcatc aagcagctcc aggcaagagt cctggctctg 1740  
gaaagatacc taagggatca acagctcatg ggaatttggg gttgctctgg aaaactcatt 1800  
tgcaccactt ctgtgccttg gaatgttagt tggagtaata aatctgtgga tgatatttgg 1860  
aataacatga cctggatgga gtgggaaaga gaaattgaca attacacaga ctatatatat 1920  
gacttacttg aaaaatcgca aaccaacaa gaaaagaatg aaaaagaatt attggaattg 1980  
gataaatggg caagtttgtg gaattggttt gacataacaa actggctgtg gtatataaga 2040  
ttattcataa t 2051

<210> 4  
 <211> 667  
 <212> PRT  
 <213> Human immunodeficiency virus type 1  
  
 <220>  
 <223> human immunodeficiency virus-1 (HIV-1) envelope  
 glycoprotein gp140 truncated version of gp 160  
 from primary R5X4 isolate 89.6  
  
 <400> 4  
 Met Arg Val Lys Glu Ile Arg Lys Asn Trp Gln His Leu Arg Gly Gly  
 1 5 10 15  
 Ile Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Lys Glu Lys  
 20 25 30  
 Thr Trp Val Thr Ile Tyr Tyr Gly Val Pro Val Trp Arg Glu Ala Thr  
 35 40 45  
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val  
 50 55 60  
 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
 65 70 75 80  
 Gln Glu Val Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
 85 90 95  
 Asn Asn Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp  
 100 105 110  
 Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
 115 120 125  
 Asn Cys Thr Asn Leu Asn Ile Thr Lys Asn Thr Thr Asn Pro Thr Ser  
 130 135 140  
 Ser Ser Trp Gly Met Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe  
 145 150 155 160  
 Tyr Ile Thr Thr Ser Ile Arg Asn Lys Val Lys Lys Glu Tyr Ala Leu  
 165 170 175  
 Phe Asn Arg Leu Asp Val Val Pro Ile Glu Asn Thr Asn Asn Thr Lys  
 180 185 190  
 Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro  
 195 200 205  
 Lys Val Ser Phe Gln Pro Ile Pro Ile His Tyr Cys Val Pro Ala Gly  
 210 215 220  
 Phe Ala Met Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Ser Gly Pro  
 225 230 235 240  
 Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val  
 245 250 255  
 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Asp Ile  
 260 265 270

Val	Ile	Arg	Ser	Glu	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr	Ile	Ile	Val	275	280	285
Gln	Leu	Asn	Glu	Ser	Val	Val	Ile	Asn	Cys	Thr	Arg	Pro	Asn	Asn	Asn	290	295	300
Thr	Arg	Arg	Arg	Leu	Ser	Ile	Gly	Pro	Gly	Arg	Ala	Phe	Tyr	Ala	Arg	305	310	315
Arg	Asn	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Arg	325	330	335
Ala	Lys	Trp	Asn	Asn	Thr	Leu	Gln	Gln	Ile	Val	Ile	Lys	Leu	Arg	Glu	340	345	350
Lys	Phe	Arg	Asn	Lys	Thr	Ile	Ala	Phe	Asn	Gln	Ser	Ser	Gly	Gly	Asp	355	360	365
Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu	Phe	Phe	Tyr	370	375	380
Cys	Asn	Thr	Ala	Gln	Leu	Phe	Asn	Ser	Thr	Trp	Asn	Val	Thr	Gly	Gly	385	390	395
Thr	Asn	Gly	Thr	Glu	Gly	Asn	Asp	Ile	Ile	Thr	Leu	Gln	Cys	Arg	Ile	405	410	415
Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Lys	Val	Gly	Lys	Ala	Met	Tyr	Ala	420	425	430
Pro	Pro	Ile	Thr	Gly	Gln	Ile	Arg	Cys	Ser	Ser	Asn	Ile	Thr	Gly	Leu	435	440	445
Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Ser	Thr	Glu	Thr	Glu	Thr	Glu	Ile	450	455	460
Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	465	470	475
Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Ile	Gly	Val	Ala	Pro	Thr	485	490	495
Arg	Ala	Lys	Arg	Arg	Thr	Val	Gln	Arg	Glu	Lys	Arg	Ala	Val	Gly	Ile	500	505	510
Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	515	520	525
Ala	Ala	Ser	Val	Thr	Leu	Thr	Val	Gln	Ala	Arg	Leu	Leu	Leu	Ser	Gly	530	535	540
Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	545	550	555
His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	565	570	575
Val	Leu	Ala	Leu	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Met	Gly	Ile	580	585	590

Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ser Val Pro Trp Asn  
 595 600 605  
 Val Ser Trp Ser Asn Lys Ser Val Asp Asp Ile Trp Asn Asn Met Thr  
 610 615 620  
 Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Asp Tyr Ile Tyr  
 625 630 635 640  
 Asp Leu Leu Glu Lys Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Glu  
 645 650 655  
 Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp  
 660 665

<210> 5  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:cleavage site  
 and start of gp41 in gp160 (env 89.6)

<400> 5  
 caaagagaaa aaagagcagt gggaata

27

<210> 6  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:cleavage site  
 and start of gp41 in gp160 (env 89.6)

<400> 6  
 Arg Glu Lys Arg Ala Val Gly Ile  
 1 5

<210> 7  
 <211> 667  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:gp140-26  
 (soluble, secreted protein)

<400> 7  
 Lys Glu Lys Thr Trp Val Thr Ile Tyr Tyr Gly Val Pro Val Trp Arg  
 1 5 10 15  
 Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp  
 20 25 30  
 Thr Glu Val His Asn Gly Trp Ala Thr His Ala Cys Val Ala Thr Asp  
 35 40 45



Pro	Asn	Pro	Gln	Glu	Val	Val	Leu	Gly	Asn	Val	Thr	Glu	Asn	Phe	Asn		
	50					55					60						
Met	Trp	Lys	Asn	Asn	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Ser		
	65				70					75					80		
Leu	Leu	Asp	Glu	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys		
				85					90					95			
Val	Thr	Leu	Asn	Cys	Thr	Asn	Leu	Asn	Ile	Thr	Lys	Asn	Thr	Thr	Asn		
			100					105					110				
Pro	Thr	Ser	Ser	Ser	Leu	Gly	Met	Met	Glu	Lys	Gly	Glu	Ile	Lys	Asn		
		115					120					125					
Cys	Ser	Phe	Tyr	Ile	Thr	Thr	Ser	Ile	Arg	Asn	Lys	Val	Lys	Lys	Glu		
	130					135					140						
Tyr	Ala	Leu	Phe	Asn	Arg	Leu	Asp	Val	Val	Pro	Ile	Glu	Asn	Thr	Asn		
	145				150					155					160		
Asn	Thr	Lys	Tyr	Arg	Leu	Ile	Ser	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln		
				165					170						175		
Ala	Cys	Pro	Lys	Val	Phe	Phe	Gln	Pro	Ile	Ala	Ile	His	Tyr	Cys	Val		
			180					185					190				
Pro	Ala	Gly	Phe	Ala	Met	Leu	Lys	Cys	Asn	Asn	Lys	Thr	Phe	Asn	Gly		
		195					200					205					
Ser	Gly	Pro	Cys	Thr	Asn	Val	Ser	Thr	Val	Pro	Cys	Thr	His	Gly	Ile		
	210					215					220						
Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu		
	225				230					235					240		
Glu	Asp	Ile	Val	Ile	Arg	Ser	Gly	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr		
				245					250					255			
Ile	Ile	Val	Gln	Leu	Asn	Glu	Ser	Val	Val	Ile	Asn	Cys	Thr	Arg	Pro		
			260					265					270				
Asn	Asn	Asn	Thr	Arg	Arg	Arg	Leu	Ser	Ile	Gly	Pro	Gly	Arg	Ala	Phe		
		275					280					285					
Tyr	Ala	Arg	Arg	Asn	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn		
	290					295					300						
Ile	Ser	Arg	Ala	Lys	Leu	Asn	Asn	Thr	Leu	Gln	Gln	Ile	Val	Ile	Lys		
	305				310					315					320		
Leu	Arg	Glu	Lys	Phe	Arg	Asn	Lys	Thr	Ile	Ala	Phe	Asn	Gln	Ser	Ser		
				325					330					335			
Gly	Gly	Asp	Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu		
			340					345					350				
Phe	Phe	Tyr	Cys	Asn	Thr	Ala	Gln	Leu	Phe	Asn	Ser	Thr	Leu	Asn	Val		
		355					360					365					

Thr	Gly	Gly	Thr	Asn	Gly	Thr	Glu	Glu	Asn	Asp	Ile	Ile	Thr	Leu	Gln	370	375	380
Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Lys	Val	Gly	Lys	Ala	385	390	395
Met	Tyr	Ala	Pro	Pro	Ile	Thr	Gly	Gln	Ile	Ile	Cys	Ser	Ser	Asn	Ile	405	410	415
Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Ser	Thr	Glu	Thr	Glu	420	425	430
Thr	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	435	440	445
Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Ile	Gly	Val	450	455	460
Ala	Pro	Thr	Arg	Ala	Lys	Arg	Arg	Thr	Cys	Gln	Gly	Gly	Ile	Asp	Gly	465	470	475
Ile	Leu	Gln	Ile	Ser	Gly	Ser	Gly	Ser	Gly	Gly	Ser	Gly	Gln	Gly	Ser	485	490	495
Ser	Ser	Gly	Gly	Ala	Gly	Gly	Lys	Gly	Ala	Val	Gly	Ile	Gly	Ala	Val	500	505	510
Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Arg	Ser	515	520	525
Val	Thr	Leu	Thr	Val	Gln	Ala	Arg	Leu	Leu	Leu	Ser	Gly	Ile	Val	Gln	530	535	540
Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Met	Leu	545	550	555
Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	565	570	575
Leu	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Met	Gly	Ile	Trp	Gly	Cys	580	585	590
Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ser	Val	Pro	Trp	Asn	Val	Ser	Trp	595	600	605
Ser	Asn	Lys	Ser	Val	Asp	Asp	Ile	Trp	Asn	Asn	Met	Thr	Trp	Met	Glu	610	615	620
Leu	Glu	Arg	Glu	Ile	Asp	Asn	Tyr	Thr	Asp	Tyr	Ile	Tyr	Asp	Leu	Leu	625	630	635
Glu	Lys	Ser	Gln	Thr	Gln	Gln	Glu	Lys	Asn	Glu	Lys	Glu	Leu	Leu	Glu	645	650	655
Leu	Asp	Lys	Trp	Ala	Ser	Leu	Trp	Lys	Leu	Val						660	665	

<210> 8  
 <211> 656  
 <212> PRT  
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:gp140-15  
 (soluble, secreted protein)

<400> 8

Lys	Glu	Lys	Thr	Trp	Val	Thr	Ile	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Arg	1	5	10	15
Glu	Ala	Thr	Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Asp	20	25	30	
Thr	Glu	Val	His	Asn	Gly	Trp	Ala	Thr	His	Ala	Cys	Val	Ala	Thr	Asp	35	40	45	
Pro	Asn	Pro	Gln	Glu	Val	Val	Leu	Gly	Asn	Val	Thr	Glu	Asn	Phe	Asn	50	55	60	
Met	Trp	Lys	Asn	Asn	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	65	70	75	80
Leu	Leu	Asp	Glu	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	85	90	95	
Val	Thr	Leu	Asn	Cys	Thr	Asn	Leu	Asn	Ile	Thr	Lys	Asn	Thr	Thr	Asn	100	105	110	
Pro	Thr	Ser	Ser	Ser	Leu	Gly	Met	Met	Glu	Lys	Gly	Glu	Ile	Lys	Asn	115	120	125	
Cys	Ser	Phe	Tyr	Ile	Thr	Thr	Ser	Ile	Arg	Asn	Lys	Val	Lys	Lys	Glu	130	135	140	
Tyr	Ala	Leu	Phe	Asn	Arg	Leu	Asp	Val	Val	Pro	Ile	Glu	Asn	Thr	Asn	145	150	155	160
Asn	Thr	Lys	Tyr	Arg	Leu	Ile	Ser	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln	165	170	175	
Ala	Cys	Pro	Lys	Val	Phe	Phe	Gln	Pro	Ile	Ala	Ile	His	Tyr	Cys	Val	180	185	190	
Pro	Ala	Gly	Phe	Ala	Met	Leu	Lys	Cys	Asn	Asn	Lys	Thr	Phe	Asn	Gly	195	200	205	
Ser	Gly	Pro	Cys	Thr	Asn	Val	Ser	Thr	Val	Pro	Cys	Thr	His	Gly	Ile	210	215	220	
Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	225	230	235	240
Glu	Asp	Ile	Val	Ile	Arg	Ser	Gly	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr	245	250	255	
Ile	Ile	Val	Gln	Leu	Asn	Glu	Ser	Val	Val	Ile	Asn	Cys	Thr	Arg	Pro	260	265	270	

Asn	Asn	Asn	Thr	Arg	Arg	Arg	Leu	Ser	Ile	Gly	Pro	Gly	Arg	Ala	Phe	275	280	285
Tyr	Ala	Arg	Arg	Asn	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	290	295	300
Ile	Ser	Arg	Ala	Lys	Leu	Asn	Asn	Thr	Leu	Gln	Gln	Ile	Val	Ile	Lys	305	310	315
Leu	Arg	Glu	Lys	Phe	Arg	Asn	Lys	Thr	Ile	Ala	Phe	Asn	Gln	Ser	Ser	325	330	335
Gly	Gly	Asp	Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu	340	345	350
Phe	Phe	Tyr	Cys	Asn	Thr	Ala	Gln	Leu	Phe	Asn	Ser	Thr	Leu	Asn	Val	355	360	365
Thr	Gly	Gly	Thr	Asn	Gly	Thr	Glu	Glu	Asn	Asp	Ile	Ile	Thr	Leu	Gln	370	375	380
Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Lys	Val	Gly	Lys	Ala	385	390	395
Met	Tyr	Ala	Pro	Pro	Ile	Thr	Gly	Gln	Ile	Ile	Cys	Ser	Ser	Asn	Ile	405	410	415
Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Ser	Thr	Glu	Thr	Glu	420	425	430
Thr	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	435	440	445
Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Ile	Gly	Val	450	455	460
Ala	Pro	Thr	Arg	Ala	Lys	Arg	Arg	Thr	Cys	Gln	Gly	Gly	Ile	Asp	Gly	465	470	475
Ile	Leu	Gln	Ile	Ser	Ser	Ser	Gly	Gly	Ala	Gly	Gly	Lys	Gly	Ala	Val	485	490	495
Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	500	505	510
Met	Gly	Ala	Arg	Ser	Val	Thr	Leu	Thr	Val	Gln	Ala	Arg	Leu	Leu	Leu	515	520	525
Ser	Gly	Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	530	535	540
Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	545	550	555
Ala	Arg	Val	Leu	Ala	Leu	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Met	565	570	575
Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ser	Val	Pro	580	585	590

Trp Asn Val Ser Trp Ser Asn Lys Ser Val Asp Asp Ile Trp Asn Asn  
595 600 605

Met Thr Trp Met Glu Leu Glu Arg Glu Ile Asp Asn Tyr Thr Asp Tyr  
610 615 620

Ile Tyr Asp Leu Leu Glu Lys Ser Gln Thr Gln Gln Glu Lys Asn Glu  
625 630 635 640

Lys Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Lys Leu Val  
645 650 655

<210> 9  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:peptide linker

<400> 9  
Gly Ile Leu Ile  
1

<210> 10  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:preferred  
peptide linker

<400> 10  
Gly Gly Ile Asp Gly Ile Leu Gln Ile Ser Ser Ser Gly Gly Ala  
1 5 10 15

<210> 11  
<211> 26  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:preferred  
peptide linker

<400> 11  
Gly Gly Ile Asp Gly Ile Leu Gln Ile Ser Gly Ser Gly Ser Gly Gly  
1 5 10 15

Ser Gly Gln Gly Ser Ser Ser Gly Gly Ala  
20 25

<210> 12  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:preferred  
         peptide linker

<400> 12  
 Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly  
     1                    5                    10                    15

<210> 13  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:preferred  
         peptide linker

<400> 13  
 Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly  
     1                    5                    10                    15

Gly Ser Gly Gly  
                 20

<210> 14  
 <211> 25  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:preferred  
         peptide linker

<400> 14  
 Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly  
     1                    5                    10                    15

Gly Ser Gly Gly Gly Gly Ser Gly Gly  
     \                20                    25

<210> 15  
 <211> 24  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
         linker

<220>  
 <221> MOD\_RES  
 <222> (7)..(24)  
 <223> amino acids at positions 7-8, 9-10, 11-12, 13-14,  
         15-16, 17-18, 19-20, 21-22, and 23-24 may be  
         present or absent

<400> 15  
Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser  
1 5 10 15

Gly Ser Gly Ser Gly Ser Gly Ser  
20

<210> 16  
<211> 48  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<220>  
<221> MOD\_RES  
<222> (13)..(48)  
<223> amino acids at positions 13-16, 17-21, 22-24,  
25-28, 29-32, 33-36, 37-40, 41-44, and 45-48 may  
be present or absent

<400> 16  
Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
20 25 30

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
35 40 45

<210> 17  
<211> 60  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<220>  
<221> MOD\_RES  
<222> (16)..(60)  
<223> amino acids at positions 16-20, 21-25, 26-30,  
31-35, 36-40, 41-45, 46-50, 51-55 and 55-60 may be  
present or absent

<400> 17  
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly  
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly  
20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly  
35 40 45

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
50 55 60

<210> 18  
<211> 72  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<220>  
<221> MOD\_RES  
<222> (19)..(72)  
<223> amino acids at positions 19-24, 25-30, 31-36,  
37-42, 43-48, 49-54, 55-60, 61-66 and 67-72 may be  
present or absent

<400> 18  
Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly  
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser  
35 40 45

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
50 55 60

Gly Ser Gly Gly Gly Gly Gly Ser  
65 70

<210> 19  
<211> 12  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<400> 19  
Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
1 5 10

<210> 20  
<211> 20  
<212> PRT  
<213> Artificial Sequence



<220>  
 <223> Description of Artificial Sequence:exemplary  
       linker

<400> 20  
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly  
       1                      5                      10                      15  
  
 Gly Gly Gly Ser  
                   20

<210> 21  
 <211> 30  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
       linker

<400> 21  
 Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
       1                      5                      10                      15  
  
 Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser  
                   20                      25                      30

<210> 22  
 <211> 42  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
       linker

<400> 22  
 Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Ser Gly Gly  
       1                      5                      10                      15  
  
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
                   20                      25                      30  
  
 Gly Gly Ser Gly Gly Gly Gly Gly Ser  
           35                      40

<210> 23  
 <211> 56  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
       linker

<400> 23  
 Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Ser  
       1                      5                      10                      15

Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Ser  
20 25 30

Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Ser  
35 40 45

Gly Gly Gly Gly Gly Gly Gly Ser  
50 55

<210> 24  
<211> 72  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<400> 24  
Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly  
1 5 10 15

Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly  
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly  
35 40 45

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly  
50 55 60

Gly Gly Gly Gly Gly Gly Gly Ser  
65 70

<210> 25  
<211> 90  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<400> 25  
Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly  
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly  
20 25 30

Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly  
35 40 45

Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
50 55 60

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
65 70 75 80

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
85 90

<210> 26  
<211> 110  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<400> 26  
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly  
1 5 10 15  
Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
20 25 30  
Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
35 40 45  
Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
50 55 60  
Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly  
65 70 75 80  
Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly  
85 90 95  
Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
100 105 110

<210> 27  
<211> 132  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:exemplary  
linker

<400> 27  
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15  
Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly  
20 25 30  
Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
35 40 45  
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
50 55 60

Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly  
 65 70 75 80  
 Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
 85 90 95  
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly  
 100 105 110  
 Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly  
 115 120 125  
 Gly Gly Gly Ser  
 130

<210> 28  
 <211> 156  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
 linker

<400> 28  
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly  
 1 5 10 15  
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly  
 20 25 30  
 Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
 35 40 45  
 Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
 50 55 60  
 Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly  
 65 70 75 80  
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly  
 85 90 95  
 Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
 100 105 110  
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
 115 120 125  
 Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly  
 130 135 140  
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser  
 145 150 155

<210> 29  
 <211> 36  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:exemplary  
 linker

<220>  
 <221> MOD\_RES  
 <222> (10)..(36)  
 <223> amino acids at positions 10-12, 13-15, 16-18,  
 19-21, 22-24, 25-27, 28-30, 31-33 and 34-36 may be  
 present or absent

<400> 29  
 Ala Gly Ser Ala Gly Ser Ala Gly Ser Ala Gly Ser Ala Gly Ser Ala  
 1 5 10 15  
 Gly Ser Ala Gly Ser Ala Gly Ser Ala Gly Ser Ala Gly Ser Ala Gly  
 20 25 30  
 Ser Ala Gly Ser  
 35

<210> 30  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:post  
 translational cleavage site

<400> 30  
 Arg Glu Lys Arg  
 1

<210> 31  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:mutated post  
 translational cleavage site

<400> 31  
 Arg Glu Ile Asp  
 1

<210> 32  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:short fragment  
 following mutated cleavage site created by  
 introduction of restriction sites

<400> 32  
Glu Phe Ile Ser  
1

<210> 33  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:polypeptide  
linker end

<400> 33  
Gly Gly Ser Gly Gly  
1 5